

Improving Returns from the Markowitz Model using GA- An Empirical Validation on the BSE

Binoy B. Nair, S. Karthik Krishna, Samridhi Singhi, Ram Prasath V.R, Muthu Arvind L.,
Abhimanyu A. and V.P Mohandas

Amrita Vishwa Vidyapeetham, Amrita Nagar, Coimbatore-Tamilnadu, India-641112.

Email: b_binoy@cb.amrita.edu

Email: skarthikkkrishna92@outlook.com, {samridhi.0302, ramprasath2406, arvindmla31, callabimanyu}@gmail.com,
vp_mohandas@amrita.edu

Abstract—Portfolio optimization is the task of allocating the investors capital among different assets in such a way that the returns are maximized while at the same time, the risk is minimized. The traditional model followed for portfolio optimization is the Markowitz model [1], [2],[3]. Markowitz model, considering the ideal case of linear constraints, can be solved using quadratic programming, however, in real-life scenario, the presence of nonlinear constraints such as limits on the number of assets in the portfolio, the constraints on budgetary allocation to each asset class, transaction costs and limits to the maximum weightage that can be assigned to each asset in the portfolio etc., this problem becomes increasingly computationally difficult to solve, ie NP-hard. Hence, soft computing based approaches seem best suited for solving such a problem. An attempt has been made in this study to use soft computing technique (specifically, Genetic Algorithms), to overcome this issue. In this study, Genetic Algorithm (GA) has been used to optimize the parameters of the Markowitz model such that overall portfolio returns are maximized with the standard deviation of the returns being minimized at the same time. The proposed system is validated by testing its ability to generate optimal stock portfolios with high returns and low standard deviations with the assets drawn from the stocks traded on the Bombay Stock Exchange (BSE). Results show that the proposed system is able to generate much better portfolios when compared to the traditional Markowitz model.

Index Terms—Genetic Algorithm, portfolio optimization, Markowitz, BSE

I. INTRODUCTION

Portfolio optimization is the process of choosing the proportions of various stocks to be held in an investment portfolio, so as to make the portfolio better than any other on the basis of some criterion. The criterion will combine, directly or indirectly, considerations of the expected value of the portfolio's rate of return as well as of the return's dispersion and possibly other measures of financial risk.

Traditional Portfolio Optimization techniques use the Markowitz model, which assists in the selection of the most efficient portfolio by analyzing various possible portfolios of the given securities. By choosing securities that do not “move” exactly together, the Markowitz model shows a method of risk reduction.

It is also called the mean-variance model because it is based on expected returns (means) and the standard deviation (Variance) of the various portfolios. The model construction is based on the following assumptions

1. Risk of a portfolio is based on variability of the returns from the portfolio.
2. Analysis is based on single period model of investment.

DOI: 02.AETACS.2013.4.97

© Association of Computer Electronics and Electrical Engineers, 2013

3. An investor either maximizes his portfolio returns for a given level of risk or maximizes his returns for the minimum risk.

To choose the best portfolio from a number of possible portfolios, each with different returns and risk, two separate decisions are to be made:

1. Determination of a set of efficient portfolios.
2. Selection of the best portfolio from the efficient set.

Markowitz model, in the ideal case, can be solved with the help of quadratic programming. However, with the addition of constraints such as restrictions on the number of assets, transaction costs, limits on the budgetary allocation and capital allocation constraints, the task of identifying solutions to the optimization problem becomes increasingly difficult. Hence, soft computing based methods have to be relied upon to obtain optimal solutions.

There have been several studies on the application of soft computing techniques to portfolio optimization. In [4], a simulated annealing based approach to solving issues in portfolio selection was presented. It was observed that the tailoring work required to account for different classes of constraints and fine-tune the parameters of the algorithm was rather delicate (especially for trading constraints. In [5], fuzzy set theory and GA was used to generate optimal portfolios with assets selected from the Istanbul Stock Exchange, however, it is seen that a certain degree of domain expertise is required for generating the fuzzy membership functions. In [6], a variant of Hopfield network was used to generate optimal portfolio and the results were compared to those obtained using other approaches such as GA, tabu search and simulated annealing. In [7], application of GA to generate optimal portfolios was presented. In [8], a survey of applications of different evolutionary algorithms for solving portfolio optimization was presented.

II. MARKOWITZ MODEL FOR PORTFOLIO OPTIMIZATION

With any set of assets and their corresponding weights and individual returns, we can find the expected return by finding the weighted average of the individual returns of all the assets as:

$$E(\text{ret}_p) = \sum_{k=1}^N w_k E(\text{ret}_k) \quad (1)$$

where:

$$\sum_{k=1}^N w_k = 1$$

N = the number of assets

w_k = the weight of asset k

$\text{ret}_{k,p}$ = the rate of return of asset k in portfolio p

$E()$ = the expectation function

Risk associated with a portfolio is typically measured by the variance. The variance of a single asset is the expected value of the sum of the squared deviations from the mean. The variance of a portfolio combination of assets is equal to the weighted average covariance of the returns on its individual securities:

$$\text{Variance}(\text{ret}_p) = \sigma_p^2 = \sum_{k=1}^N \sum_{l=1}^N w_k w_l \text{CoVar}(\text{ret}_k, \text{ret}_l) \quad (2)$$

Covariance in terms of correlation coefficient is:

$$\text{CoVar}(\text{ret}_k, \text{ret}_l) = \rho_{kl} \sigma_k \sigma_l = \sigma_{kl} \quad (3)$$

Where,

ρ_{kl} = correlation coefficient between ret_k and ret_l

ret_i = the rate of return of asset i

σ_i = standard deviation of ret_i

The variance of the portfolio p is now given as:

$$\text{Variance}(\text{ret}_p) = \sum_{k=1}^N \sum_{l=1}^N w_k w_l \sigma_{kl} \quad (4)$$

The standard deviation of the portfolio is the square root of the variance. The mean return of every asset is its average value in the sample space. The variance is the average of the squared deviations of the sampled average and covariance gives the average value of product of deviations.

Suppose a portfolio p is constructed of two assets A_1 and A_2 with their weights in the portfolio being represented by w_{A_1} and w_{A_2} respectively. Also $w_{A_1} + w_{A_2} = 1$. The expected return of the portfolio p will be:

$$E(\text{ret}_p) = w_{A_1}E(\text{ret}_{A_1}) + w_{A_2}E(\text{ret}_{A_2}) \quad (5)$$

The variance of the rate of return on portfolio p is

$$\text{Variance}(\text{ret}_p) = \sigma_p^2 = (w_{A_1}\sigma_{A_1} + w_{A_2}\sigma_{A_2})^2 \quad (6)$$

which reduces to

$$\sigma_p^2 = w_{A_1}^2\sigma_{A_1}^2 + w_{A_2}^2\sigma_{A_2}^2 + 2w_{A_1}w_{A_2}\rho_{A_1A_2}\sigma_{A_1}\sigma_{A_2} \quad (7)$$

The correlation between the asset components (ρ) is inversely proportional to portfolio risk. That is, the lower (even negative) the correlation, the lower the portfolio risk. For example let's assume that $\rho_{A_1A_2} = 1$. Then, the right-hand side of equation (7) is a perfect square and simplifies to $\sigma_p = w_{A_1}\sigma_{A_1} + w_{A_2}\sigma_{A_2}$.

In this special case, the portfolio mean and standard deviations are simple weighted averages of the asset returns.

A plot between the (mean) returns of the portfolio on the y-axis and the standard deviation from the mean (also called the risk factor) on the x-axis, was proposed by Markowitz to identify the portfolio with maximum returns or minimum risks. All asset combinations, resulting in different possible portfolios, are plotted in this risk-return space. The set of portfolios that give the most optimal returns form the efficient frontier.

For a given appetite of risk the investor has, the portfolio lying on the efficient frontier gives the best portfolio the investor can have. The portfolio selection problem can be formulated as a quadratic program as follows:

$$\min w^T Q w \quad (8)$$

with

$$\begin{aligned} \sum_{k=1}^N w_k \text{ret}_k &\geq \text{ret}_p \\ \sum_{k=1}^N w_k &= 1, \\ w_k &\geq 0 \quad \forall k \text{ in } [1, N] \text{ and,} \\ Pw &\leq q \quad (\text{additional constraints}) \end{aligned}$$

where,

$$\begin{aligned} Q &\in \mathbb{R}^{N \times N} \\ q &\in \mathbb{R}^M \\ P &\in \mathbb{R}^{N \times M} \end{aligned}$$

The above formulated quadratic problem is convex and can be solved using quadratic programming. However, presence of constraints such as upper and lower bounds on the asset allocation, upper and lower limits on budgetary constraints and the variation in the number of assets (N) cause additional constraints and require modification of the constraints specified in (8). This results in an increase in complexity of the optimization problem and makes it a good candidate for application of soft computing based techniques such as GA.

III. GENETIC ALGORITHMS

GA is a parallel search algorithm that can be used for optimizing highly nonlinear objective functions. The algorithm starts with a set of candidate solutions randomly selected from the specified search space. The collection of these candidate solutions is called the population. The best candidate solutions from the population are selected based on a measure of fitness. The selected solutions are called the parents. The selected parents undergo crossover and mutation to yield a new set of candidate solutions, called children. The collection of parents and children together form the mating pool, which form the basis for the next 'generation' of solutions. The process of selection, crossover and mutation is repeated every generation and this results in improvement of the fitness of the population over generations. The GA terminates when either

a predefined terminating condition is encountered or when there is no further improvement in the fitness of the population over many generations.

In the present study, the variables which need to be optimized using GA include: the upper and lower bounds on capital allocation to an individual asset, the upper and lower bounds on the budget that can be allocated to assets with the rest being held in cash, and the number of assets to be selected to make up the portfolio.

IV. SYSTEM SPECIFICATIONS

The GA optimized Markowitz model is validated on the assets drawn from the Bombay Stock Exchange (BSE). A total of 30 large market capitalization companies stocks (those belonging to Group-A of the BSE) were selected initially. The time frame of the data was from 22/10/2010 to 26/07/2013, collected on daily basis (daily closing prices). Three different categories of portfolios were created, with portfolio category 1 using GA to optimize only the upper and lower bounds on asset allocation, portfolio category 2 using GA to optimize the upper, lower bounds on asset allocation and upper, lower bounds on budgetary allocation and portfolio category 3 using GA to optimize upper, lower bounds on asset allocation, upper, lower bounds on budgetary allocation as well as the number of assets to be considered in the portfolio. Three different proxies for the risk free returns, investments in which, is available to the general public, were also considered.

The system details are given as follows:

A. Assets

A total of 30 stocks belonging to group A [9], listed on the Bombay Stock Exchange are considered. Some or all of these stocks could constitute the final optimized portfolio. The stocks considered are: TCS, Reliance Inds., ITC, ONGC, Coal India, UltraTech Cement, HDFC Bank, Hindustan Unilever, BhartiAirtel, Wipro, NTPC, SBI, ICICI Bank, Tata Motors, HCL Technologies, Sun Pharma, Bajaj Auto, Indian Oil, Idea Cellular, Kotak Mahindra Bank, Larsen and Tubro, NMDC, Mahindra & Mahindra, Nestle India, Axis Bank Ltd., Asian Paints, Hero Moto Corp, CIPLA, Maruti Suzuki India, GAIL(India).

B. Risk free rate of return

Comparative study was carried out based on three different proxies for the risk-free rate:

1. Government of India bonds[10]: returns 8% p.a
2. Post office time deposit[11]: 8.5% returns p.a
3. National Savings Certificate (NSC) IX Issue (10 years)[12]: 8.8% returns p.a

C. Buying and selling costs

The buying and selling costs are set to 0.5%. This figure was arrived at from the information available with different brokerage houses. There are varying transaction costs for different plans offered by each of these houses and there does not appear to be any standardization in existence as far as transaction costs are considered. The transaction costs could be as low as 0.1% (subject to conditions) [13] selecting which, would have further boosted the returns generated by our optimized portfolios but to err on the side of caution, a transaction cost of 0.5% [14] was considered. Also, no short selling is to be allowed.

D. GA parameters

The objective function chosen for the purpose of optimization was:

$\text{minimize } (1/(1+\text{avg portfolio returns}(X)))$

Where,

$X = \{\text{upper bound, lower bound}\}$ for portfolio category 1,

$X = \{\text{upper bound, lower bound, upper budget, lower budget}\}$ for portfolio category 2, and

$X = \{\text{upper bound, lower bound, upper budget, lower budget, number of assets}\}$ for portfolio category 3.

Following GA parameters were used for the optimization process:

Population size: 20

Max generations: 100

Selection: Stochastic uniform (involves laying out a line with each parent corresponding to a section of the line of length proportional to its scaled value and then making a pointer move along the line in equal sized steps. At every step, the parent from the section where the pointer now lies, is selected.

Reproduction: The crossover operation is carried out using a binary crossover mask. A random binary vector (called the mask) of the same length as the two parents is created. In The positions where 1's are seen in the mask, the gene is taken from the first parent and the positions in the mask where 0's appear, gene is taken from the second parent to finally yield the child. 80% of the population in the next generation is created using this crossover. Remaining are elite children that are the best fit members of the current population and survive to the next generation.

Mutation: A Gaussian method of mutation is followed where a random number from a Gaussian distribution with mean 0 is added to each entry of the parent vector.

V. RESULTS

Performance of the proposed system is presented in the form of three tables Table I-III. The average daily returns produced by the portfolio on the efficient frontier with maximum returns, is presented. Efficient frontiers obtained with the use of default constraints and those obtained with the five GA-optimized parameters are presented in Fig. 1-3. Parameters which are not optimized (in portfolio categories 1 and 2) are set to their ideal valuesie. no upper and lower bounds on asset allocation, no cash holdings at any time and the entire set of 30 assets to be considered. Transaction costs are kept at 0.5% for all the three categories

TABLE I. THREE CATEGORIES OF GA OPTIMIZED PORTFOLIOS WITH RISK FREE RETURNS AS GOI BOND RETURNS

Portfolio category	GA optimized parameters					Avg .Daily Returns	Std Deviation of returns
	Lower Bound	Upper Bound	Lower Budget	Upper Budget	Num Assets		
1	0.02	0.8	----	----	---	-9.8406e-5	0.0083
2	0.0129	0.875	0.3203	1	---	-9.8406e-5	0.0083
3	0.0143	0.8	0.01	1	2	0.0011	0.0142

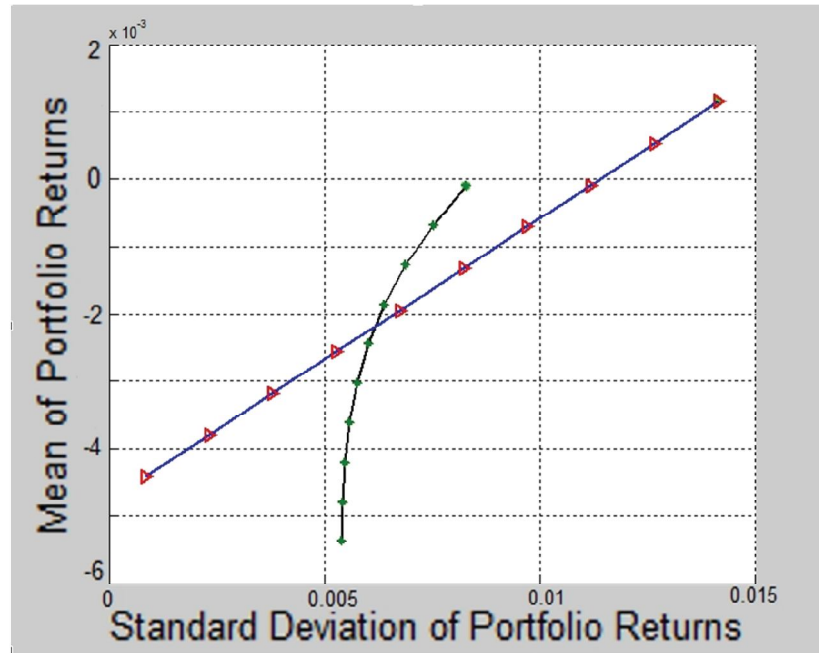


Figure 1. Efficient frontiers for GoI bonds as risk free proxy. Portfolios on the efficient frontier with ideal values are represented using '*' and portfolios on the efficient frontier for GA optimized parameters are represented using 'Δ'

TABLE II.THREE CATEGORIES OF GA OPTIMIZED PORTFOLIOS WITH RISK FREE RETURNS AS POST OFFICE TIME DEPOSITS RETURNS

Portfolio category	GA optimized parameters					Avg .Daily Returns	StdDeviation of returns
	Lower Bound	Upper Bound	Lower Budget	Upper Budget	Num Assets		
1	0.020	0.8	---	---	---	-9.8406e-5	0.0083
2	0.01	0.8	0.02368	1	---	-9.8406e-5	0.0083
3	0.01	0.8	0.06	1	2	0.0011	0.0142

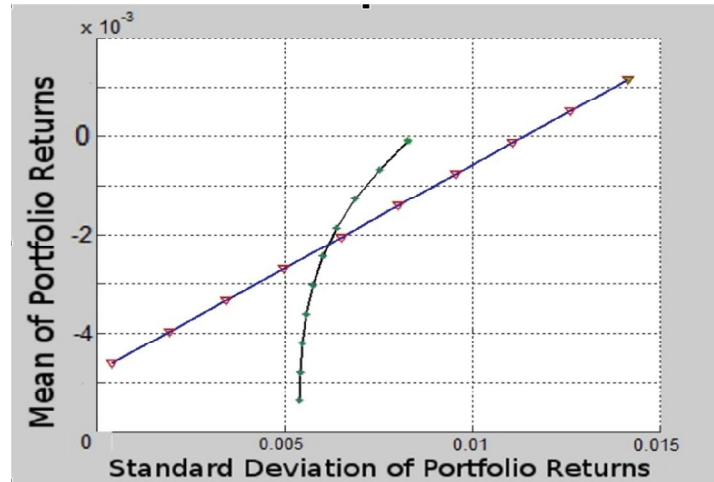


Figure 2. Efficient frontiers for post office time deposits as risk free proxy. Portfolios on the efficient frontier with ideal values are represented using '*' and portfolios onthe efficient frontier for GA optimized parameters are represented using 'Δ'

TABLE III.THREE CATEGORIES OF GA OPTIMIZED PORTFOLIOS WITH RISK FREE RETURNS AS NSC RETURNS

Portfolio category	GA optimized parameters					Avg .Daily Returns	Std deviation of returns
	Lower Bound	Upper Bound	Lower Budget	Upper Budget	Num Assets		
1	0.0200	0.8	---	---	---	-9.8406e-5	0.0083
2	0.0165	0.8241	0.2977	1	---	-9.8406e-5	0.0083
3	0.0158	0.8333	0.7112	1	2	0.0011	0.0142

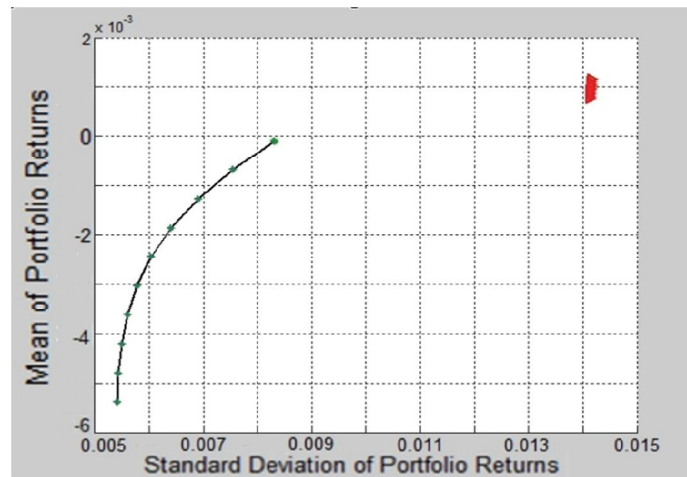


Figure 3. Efficient frontiers for post office time deposits as risk free proxy. Portfolios on the efficient frontier with ideal values are represented using '*' and portfolios onthe efficient frontier for GA optimized parameters are represented using 'Δ'

VI. CONCLUSIONS

It can be seen from the results that the use of GA to optimize the Markowitz model parameters significantly improves the returns of the portfolio. Optimizing the number of assets in the portfolio has a significant impact on the returns generated. It was also seen that there is no significant change in the returns generated by the three categories of portfolios for the three different proxies of risk free returns. This could be due to the fact that the daily returns generated by these proxies are not very different from one another. When compared to the returns generated by investing the capital in BSE-Sensex (the representative index for the stocks traded on the BSE) using the traditional buy-and-hold strategy over the same time frame of 710 days, the returns of the three portfolio categories is presented in Fig. 4.

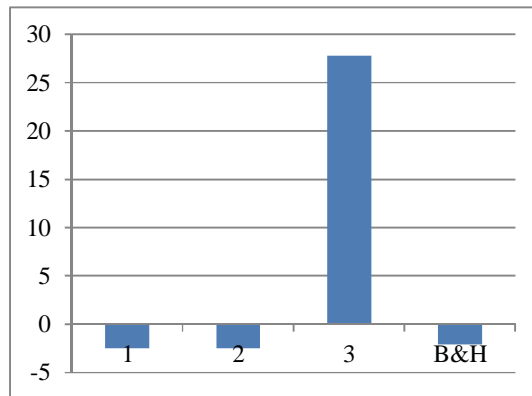


Figure 4 Percentage returns of the three categories of portfolios (1, 2 and 3) compared with buy and hold returns (B&H)

From Figure 4, it is clearly seen that the GA based optimization of the Markowitz model parameters significantly improves the returns when compared to the traditional buy-and-hold strategy.

REFERENCES

- [1] H. Markowitz, "Portfolio Selection", The Journal of Finance, vol. 7, no. 1, pp. 77-91, March 1952.
- [2] H. Markowitz, Portfolio Selection: Efficient Diversification of Investments, John Wiley & Sons, Inc., New York, 1959.
- [3] M.C. Steinbach, "Markowitz Revisited: Mean-Variance Models in Financial Portfolio Analysis", SIAM Review, vol. 43, no. 1, pp. 31-85.
- [4] Y. Crama, and M. Schyns, "Simulated annealing for complex portfolio selection problems", European Journal of operational research, vol. 150, no. 3, pp. 546-571, 2003.
- [5] R. H. Abiyev and M. Menekay, "Fuzzy portfolio selection using genetic algorithm", Soft Computing, vol. 11, no. 12, pp. 1157-1163, 2007.
- [6] A. Fernández and S. Gómez, "Portfolio selection using neural networks", Computers & Operations Research, vol. 34, no. 4, pp. 1177-1191, 2007.
- [7] J. Budík and R. Doskočil, "Soft Computing as A Tool to Optimize an Investment Portfolio", Intelektinė ekonomika, vol. 5, no. 3, pp. 359-370, 2011.
- [8] Omar Rifki and Hirotaka Ono, A survey of computational approaches to portfolio optimization by genetic algorithms, 2012. Available: <http://hdl.handle.net/2324/25317>.
- [9] List of Group A BSE stocks. Available online: http://www.bseindia.com/downloads/About/file/Criteria_for_Group_A.pdf
- [10] GOI Savings Bonds details. Available online: <http://www.icicibank.com/personal-banking/investments/goi-bonds/>.
- [11] Post office time deposits scheme details. Available online: <http://www.indiapost.gov.in/TimeDeposit.aspx>.
- [12] National Savings Certificate (NSC) IX Issue (10 years) scheme details. Available online: <http://www.indiapost.gov.in/NSC.aspx>
- [13] Stock Transaction charges. Available online: <http://www.rsec.co.in/charge-structure/rmax.html>
- [14] Stock Transaction charges. Available online: <http://www.hdfcsec.com/FAQ.aspx?docid=201008101155287343750>